Relevance Theory

Deirdre Wilson, Dan Sperber

(To appear in G. Ward and L. Horn (eds) *Handbook of Pragmatics*. Oxford : Blackwell.)

RELEVANCE THEORY*

DEIRDRE WILSON AND DAN SPERBER

1. Introduction

Relevance theory may be seen as an attempt to work out in detail one of Grice's central claims: that an essential feature of most human communication, both verbal and non-verbal, is the expression and recognition of intentions (Grice 1989: Essays 1-7, 14, 18; Retrospective Epilogue). In developing this claim, Grice laid the foundations for an inferential model of communication, an alternative to the classical code model. According to the code model, a communicator encodes her intended message into a signal, which is decoded by the audience using an identical copy of the code. According to the inferential model, a communicator provides evidence of her intention to convey a certain meaning, which is inferred by the audience on the basis of the evidence provided. An utterance is, of course, a linguistically coded piece of evidence, so that verbal comprehension involves an element of decoding. However, the linguistic meaning recovered by decoding is just one of the inputs to a non-demonstrative inference process which yields an interpretation of the speaker's meaning.[1]

The goal of inferential pragmatics is to explain how the hearer infers the speaker's meaning on the basis of the evidence provided. The relevance-theoretic account is based on another of Grice's central claims: that utterances automatically create expectations which guide the hearer towards the speaker's meaning. Grice described these expectations in terms of a Co-operative Principle and maxims of Quality (truthfulness), Quantity (informativeness), Relation (relevance) and Manner (clarity) which speakers are expected to observe (Grice 1961; 1989: 368-72): the interpretation a rational hearer should choose is the one that best satisfies those expectations.

Relevance theorists share Grice's intuition that utterances raise expectations of relevance, but question several other aspects of his account, including the need for a Co-operative Principle and maxims, the focus on pragmatic processes which contribute to implicatures rather than to explicit, truth-conditional content, the role of deliberate maxim violation in utterance interpretation, and the treatment of figurative utterances as deviations from a maxim or convention of truthfulness.[2] The central claim of relevance theory is that the expectations of relevance raised by an utterance are precise enough, and predictable enough, to guide the hearer towards the speaker's meaning. The aim is to explain in cognitively realistic terms what these expectations of relevance amount to, and how they might contribute to an empirically plausible account of comprehension. The theory has developed in several stages. A detailed version was published in *Relevance: Communication and Cognition* (Sperber & Wilson 1986a; 1987a,b) and updated in Sperber & Wilson 1995, 1998a, 2002; Wilson & Sperber 2002. Here, we will outline the main assumptions of the current version of the theory and discuss some of its implications for pragmatics.

2. Relevance and cognition

What sort of things may be relevant? Intuitively, relevance is a potential property not only of utterances and other observable phenomena, but of thoughts, memories and conclusions of inferences. In relevance-theoretic terms, any external stimulus or internal representation which provides an input to cognitive processes may be relevant to an individual at some time. According to relevance theory, utterances raise expectations of relevance not because speakers are expected to obey a Co-operative Principle and maxims or some other specifically communicative convention, but because the search for relevance is a basic feature of human cognition, which communicators may exploit. In this section, we will introduce the basic

cognitive notion of relevance and the Cognitive Principle of Relevance, which lay the foundation for the relevance-theoretic approach to pragmatics.

When is an input relevant? Intuitively, an input (a sight, a sound, an utterance, a memory) is relevant to an individual when it connects with background information he has available to yield conclusions that matter to him: say, by answering a question he had in mind, improving his knowledge on a certain topic, settling a doubt, confirming a suspicion, or correcting a mistaken impression. In relevance-theoretic terms, an input is relevant to an individual when its processing in a context of available assumptions yields a POSITIVE COGNITIVE EFFECT. A positive cognitive effect is a worthwhile difference to the individual's representation of the world – a true conclusion, for example. False conclusions are not worth having. They are cognitive effects, but not positive ones (Sperber & Wilson 1995: §3.1-2).[3]

The most important type of cognitive effect achieved by processing an input in a context is a CONTEXTUAL IMPLICATION, a conclusion deducible from the input and the context together, but from neither input nor context alone. For example, on seeing my train arriving, I might look at my watch, access my knowledge of the train timetable, and derive the contextual implication that my train is late (which may itself achieve relevance by combining with further contextual assumptions to yield further implications). Other types of cognitive effect include the strengthening, revision or abandonment of available assumptions. For example, the sight of my train arriving late might confirm my impression that the service is deteriorating, or make me alter my plans to do some shopping on the way to work. According to relevance theory, an input is RELEVANT to an individual when, and only when, its processing yields such positive cognitive effects.[4]

Intuitively, relevance is not just an all-or-none matter but a matter of degree. There is no shortage of potential inputs which might have at least some relevance for us, but we cannot attend

to them all. Relevance theory claims that what makes an input worth picking out from the mass of competing stimuli is not just that it is relevant, but that it is *more* relevant than any alternative input available to us at that time. Intuitively, other things being equal, the more worthwhile conclusions achieved by processing an input, the more relevant it will be. In relevance-theoretic terms, other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater its relevance will be. Thus, the sight of my train arriving one minute late may make little worthwhile difference to my representation of the world, while the sight of it arriving half an hour late may lead to a radical reorganisation of my day, and the relevance of the two inputs will vary accordingly.

What makes an input worth picking out from the mass of competing stimuli is not just the cognitive effects it achieves. In different circumstances, the same stimulus may be more or less salient, the same contextual assumptions more or less accessible, and the same cognitive effects easier or harder to derive. Intuitively, the greater the effort of perception, memory and inference required, the less rewarding the input will be to process, and hence the less deserving of our attention. In relevance-theoretic terms, other things being equal, the greater the PROCESSING EFFORT required, the less relevant the input will be. Thus, RELEVANCE may be assessed in terms of cognitive effects and processing effort:

(1) **Relevance of an input to an individual**

a. Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.

b. Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time.

Here is a brief and artificial illustration of how the relevance of alternative inputs might be compared in terms of effort and effect. Mary, who dislikes most meat and is allergic to chicken, rings her dinner party host to find out what is on the menu. He could truly tell her any of three things:

(2) We are serving meat.

(3) We are serving chicken.

(4) Either we are serving chicken or $(7^2 – 3)$ is not 46.

According to the characterisation of relevance in (1), all three utterances would be relevant to Mary, but (3) would be more relevant than either (2) or (4). It would be more relevant than (2) for reasons of cognitive effect: (3) entails (2), and therefore yields all the conclusions derivable from (2), and more besides. It would be more relevant than (4) for reasons of processing effort: although (3) and (4) are logically equivalent, and therefore yield exactly the same cognitive effects, these effects are easier to derive from (3) than from (4), which requires an additional effort of parsing and inference (in order to work out that the second disjunct is false and the first is therefore true). Thus, (3) would be the most relevant utterance to Mary, for reasons of both effort and effect. More generally, when similar amounts of effort are required, the effect factor is decisive in determining degrees of relevance, and when similar amounts of effect are achievable, the effort factor is decisive.

This characterisation of relevance is comparative rather than quantitative: it makes clear comparisons possible in some cases (e.g. (2)–(4)), but not in all. While quantitative notions of relevance might be worth exploring from a formal point of view[5], it is the comparative rather than the quantitative notion that is likely to provide the best starting point for constructing a psychologically plausible theory. In the first place, it is highly unlikely that individuals have to

compute numerical values for effort and effect when assessing relevance "from the inside". Such computation would itself be effort-consuming and therefore detract from relevance. Moreover, even when individuals are clearly capable of computing numerical values (for weight or distance, for example), they generally have access to more intuitive methods of assessment which are comparative rather than quantitative, and which are in some sense more basic. In the second place, while some aspects of human cognitive processes can already be measured "from the outside" (e.g. processing time) and others may be measurable in principle (e.g. number of contextual implications), it is quite possible that others are not measurable at all (e.g. strength of implications, level of attention). As noted in *Relevance* (124-32), it therefore seems preferable to treat effort and effect as *non-representational* dimensions of mental processes: they exist and play a role in cognition whether or not they are mentally represented; and when they are mentally represented, it is in the form of intuitive comparative judgements rather than absolute numerical ones. The same is true of relevance, which is a function of effort and effect.[6, 7]

Given the characterisation of relevance in (1), aiming to maximise the relevance of the inputs one processes is simply a matter of making the most efficient use of the available processing resources. No doubt this is something we would all want to do, given a choice. Relevance theory claims that humans do have an automatic tendency to maximise relevance, not because we have a choice in the matter – we rarely do – but because of the way our cognitive systems have evolved. As a result of constant selection pressure towards increasing efficiency, the human cognitive system has developed in such a way that our perceptual mechanisms tend automatically to pick out potentially relevant stimuli, our memory retrieval mechanisms tend automatically to activate potentially relevant assumptions, and our inferential mechanisms tend spontaneously to process them in the most productive way. Thus, while we are all likely to notice the sound of glass breaking in our vicinity, we are likely to attend to it more, and process it more deeply, when our

memory and inference mechanisms identify it as the sound of *our* glass breaking, and compute the consequences that are likely to be most worthwhile for us. This universal tendency is described in the First, or Cognitive, Principle of Relevance (Sperber & Wilson 1995: §3.1-2):

(5) **Cognitive Principle of Relevance**

Human cognition tends to be geared to the maximisation of relevance.

It is against this cognitive background that inferential communication takes place.

3. Relevance and communication

The universal cognitive tendency to maximise relevance makes it possible, at least to some extent, to predict and manipulate the mental states of others. Knowing of your tendency to pick out the most relevant stimuli in your environment and process them so as to maximise their relevance, I may be able to produce a stimulus which is likely to attract your attention, to prompt the retrieval of certain contextual assumptions and to point you towards an intended conclusion. For example, I may leave my empty glass in your line of vision, intending you to notice and conclude that I might like another drink. As Grice pointed out, this is not yet a case of inferential communication because, although I did intend to affect your thoughts in a certain way, I gave you no evidence that I had this intention. Inferential communication is not just a matter of intending to affect the thoughts of an audience; it is a matter of getting them to recognise that one has this intention. When I quietly leave my glass in your line of vision, I am not engaging in inferential communication, but merely exploiting your natural cognitive tendency to maximise relevance. Inferential communication – what relevance theory calls OSTENSIVE-INFERENTIAL COMMUNICATION for reasons that will shortly become apparent – involves an extra layer of intention:

(6) **Ostensive-inferential communication**

a. **The informative intention**:

The intention to inform an audience of something.

b. **The communicative intention**:

The intention to inform the audience of one's informative intention.[8]

Understanding is achieved when the communicative intention is fulfilled – that is, when the audience recognises the informative intention. (Whether the informative intention itself is fulfilled depends on how much the audience trusts the communicator. There is a gap between understanding and believing. For understanding to be achieved, the informative intention must be recognised, but it does not have to be fulfilled.)

How does the communicator indicate to the audience that she is trying to communicate with them in this overt, intentional way? Instead of covertly leaving my glass in your line of vision, I might touch your arm and point to my empty glass, wave it at you, ostentatiously put it down in front of you, stare at it meaningfully, or say "My glass is empty". More generally, ostensive-inferential communication involves the use of an OSTENSIVE STIMULUS, designed to attract an audience's attention and focus it on the communicator's meaning. Relevance theory claims that use of an ostensive stimulus may create precise and predictable expectations of relevance not raised by other stimuli. In this section, we will describe these expectations and show how they may help the audience to identify the communicator's meaning.

The fact that ostensive stimuli create expectations of relevance follows from the definition of an ostensive stimulus and the Cognitive Principle of Relevance. An ostensive stimulus is designed to attract the audience's attention. Given the universal tendency to maximise relevance, an audience will only pay attention to a stimulus that seems relevant enough. By producing an

ostensive stimulus, the communicator therefore encourages her audience to presume that it is relevant enough to be worth processing. This need not be a case of Gricean co-operation. Even a self-interested, deceptive or incompetent communicator manifestly intends her audience to assume that her stimulus is relevant enough to be worth processing – why else would he pay attention?[9] This is the basis for the Second, or Communicative, Principle of Relevance, which applies specifically to ostensive-inferential communication:

(7) Communicative Principle of Relevance

Every ostensive stimulus conveys a presumption of its own optimal relevance.

The Communicative Principle of Relevance and the notion of OPTIMAL RELEVANCE (see below) are the key to relevance-theoretic pragmatics.

An ostensive stimulus, then, creates a PRESUMPTION OF RELEVANCE. The notion of optimal relevance is meant to spell out what the audience of an act of ostensive communication is entitled to expect in terms of effort and effect:

(8) Optimal relevance

An ostensive stimulus is optimally relevant to an audience iff:

a. It is relevant enough to be worth the audience's processing effort;

b. It is the most relevant one compatible with communicator's abilities and preferences.

According to clause (a) of this definition of optimal relevance, the audience is entitled to expect the ostensive stimulus to be at least relevant enough to be worth processing. Given the argument of the last section that a stimulus is worth processing only if it is more relevant than any alternative input available at the time, this is not a trivial claim. Indeed, in order to satisfy the presumption of relevance conveyed by an ostensive stimulus, the audience may have to draw

stronger conclusions than would otherwise have been warranted. For example, if you just happen to notice my empty glass, you may be entitled to conclude that I *might* like a drink. If I deliberately wave it about in front of you, you would generally be justified in drawing the stronger conclusion that I *would* like a drink.

According to clause (b) of the definition of optimal relevance, the audience of an ostensive stimulus is entitled to even higher expectations than this. The communicator wants to be understood. It is therefore in her interest – within the limits of her own capabilities and preferences – to make her ostensive stimulus as easy as possible for the audience to understand, and to provide evidence not just for the cognitive effects she aims to achieve in her audience but also for further cognitive effects which, by holding his attention, will help her achieve her goal. For instance, the communicator's goal might be to inform her audience that she has begun writing her paper. It may be effective for her, in pursuit of this goal, to volunteer more specific information and say, "I have already written a third of the paper." In the circumstances, her audience would then be entitled to understand her as saying that she has she has written only a third of the paper, for if she had written two thirds (say), she would normally be expected to say so, given clause (b) of the definition of optimal relevance.

Communicators, of course, are not omniscient, and they cannot be expected to go against their own interests and preferences in producing an utterance. There may be relevant information that they are unable or unwilling to provide, and ostensive stimuli that would convey their intentions more economically, but that they are unwilling to produce, or unable to think of at the time. All this is allowed for in clause (b) of the definition of optimal relevance, which states that the ostensive stimulus is the most relevant one (i.e. yielding the greatest effects, in return for the smallest processing effort) that the communicator is WILLING AND ABLE to produce (see Sperber & Wilson 1995: §3.3 and 266-78).

This approach sheds light on some cases where a communicator withholds relevant information, and which seem to present problems for Grice. Suppose I ask you a question and you remain silent. Silence in these circumstances may or may not be an ostensive stimulus. When it is not, we would naturally take it as indicating that the addressee was unable or unwilling to answer the question. If you are clearly willing to answer, I am entitled to conclude that you are unable, and if you are clearly able to answer, I am entitled to conclude that you are unwilling. When the silence is ostensive, we would like to be able to analyse it as merely involving an extra layer of intention, and hence as COMMUNICATING – or IMPLICATING – that the addressee is unable or unwilling to answer. Given the presumption of relevance and the definition of optimal relevance in (8), this is possible in the relevance-theoretic framework.[10] In Grice's framework, by contrast, the co-operative communicator's willingness to provide any required information is taken for granted, and the parallels between ostensive and non-ostensive silences are lost. On a Gricean account, violation of the first Quantity maxim ("Make your contribution as informative as required") is invariably attributed to the communicator's INABILITY – rather than UNWILLINGNESS – to provide the required information. Unwillingness to make one's contribution 'such as is required' is a violation of the Co-operative Principle, and suspension of the Co-operative Principle should make it impossible to convey any conversational implicatures at all.[11] We have argued that, although much communication is co-operative in the sense that the communicator is willing to provide the required information, co-operation in this sense is not essential for communication, as it is for Grice (for references, see footnote 9)).

This relevance-theoretic account of cognition and communication has practical implications for pragmatics. As noted above, verbal comprehension starts with the recovery of a linguistically encoded sentence meaning, which must be contextually enriched in a variety of ways to yield a full-fledged speaker's meaning. There may be ambiguities and referential ambivalences to

resolve, ellipses to interpret, and other underdeterminacies of explicit content to deal with.[12] There may be implicatures to identify, illocutionary indeterminacies to resolve, metaphors and ironies to interpret. All this requires an appropriate set of contextual assumptions, which the hearer must also supply. The Communicative Principle of Relevance and the definition of optimal relevance suggest a practical procedure for performing these subtasks and constructing a hypothesis about the speaker's meaning. The hearer should take the linguistically encoded sentence meaning; following a path of least effort, he should enrich it at the explicit level and complement it at the implicit level until the resulting interpretation meets his expectation of relevance:

(9) Relevance-theoretic comprehension procedure

a. Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.

b. Stop when your expectations of relevance are satisfied.

Given clause (b) of the definition of optimal relevance in (8), it is reasonable for the hearer to follow a path of least effort because the speaker is expected (within the limits of her abilities and preferences) to make her utterance as easy as possible to understand. Since relevance varies inversely with effort, the very fact that an interpretation is easily accessible gives it an initial degree of plausibility. It is also reasonable for the hearer to stop at the first interpretation that satisfies his expectations of relevance, because there should never be more than one. A speaker who wants her utterance to be as easy as possible to understand should formulate it (within the limits of her abilities and preferences) so that the first interpretation to satisfy the hearer's expectation of relevance is the one she intended to convey.[13] An utterance with two apparently satisfactory competing interpretations would cause the hearer the unnecessary extra effort of

choosing between them, and the resulting interpretation (if there were one) would not satisfy clause (b) of the definition of optimal relevance.[14]

Thus, when a hearer following the path of least effort arrives at an interpretation that satisfies his expectations of relevance, in the absence of contrary evidence, this is the most plausible hypothesis about the speaker's meaning. Since comprehension is a non-demonstrative inference process, this hypothesis may well be false; but it is the best a rational hearer can do.

4. Relevance and comprehension

In many non-verbal cases (e.g. pointing to one's empty glass, failing to respond to a question), use of an ostensive stimulus merely adds an extra layer of intention recognition to a basic layer of information that the audience might have picked up anyway. In other cases (e.g. inviting someone out to a drink by pretending to raise a glass to one's lips), the communicator's behaviour provides no direct evidence for the intended conclusion, and it is only the presumption of relevance conveyed by the ostensive stimulus which encourages the audience to devote the necessary processing resources to discovering her meaning. Either way, the range of meanings that can be non-verbally conveyed is necessarily limited by the range of concepts the communicator can evoke in her audience by drawing attention to observable features of the environment (whether preexisting or produced specifically for this purpose).

In verbal communication, speakers manage to convey a very wide range of meanings despite the fact that there is no independently identifiable basic layer of information forthe hearer to pick up. What makes it possible for the hearer to recognise the speaker's informative intention is that utterances encode logical forms (conceptual representations, however fragmentary or incomplete) which the speaker has manifestly chosen to provide as input to the hearer's inferential comprehension process. As a result, verbal communication can achieve a degree of explicitness

not available in non-verbal communication (compare pointing in the direction of a table containing glasses, ashtrays, plates, etc., and saying, “My glass is empty”).

Although the decoded logical form of an utterance is an important clue to the speaker’s intentions, it is now increasingly recognised that even the explicitly communicated content of an utterance goes well beyond what is linguistically encoded.[15] Grice talked of his Co-operative Principle and maxims mainly in connection with the recovery of implicatures, and he seems to have thought of them as playing no significant role on the explicit side. His few remarks on disambiguation and reference assignment – which he saw as falling on the explicit rather than the implicit side – suggest that he thought of them as determined by sentence meaning and contextual factors alone, without reference to pragmatic principles or speakers’ intentions,[16] and many pragmatists have followed him on this. There has thus been a tendency, even in much of the recent pragmatic literature, to treat the “primary” processes involved in the recovery of explicit content as significantly different from – i.e. less inferential, or less directly dependent on speakers’ intentions or pragmatic principles than – the “secondary” processes involved in the recovery of implicatures.[17]

Relevance theory treats the identification of explicit content as equally inferential, and equally guided by the Communicative Principle of Relevance, as the recovery of implicatures. The relevance-theoretic comprehension procedure (“Follow a path of least effort in computing cognitive effects: test interpretive hypotheses in order of accessibility, and stop when your expectations of relevance are satisfied.”) applies in the same way to the resolution of linguistic underdeterminacies at both explicit and implicit levels. The hearer’s goal is to construct a hypothesis about the speaker’s meaning which satisfies the presumption of relevance conveyed by the utterance. This overall task can be broken down into a number of sub-tasks:

(10) Sub-tasks in the overall comprehension process

a. Constructing an appropriate hypothesis about explicit content (in relevance-theoretic terms, EXPLICATURES) via decoding, disambiguation, reference resolution, and other pragmatic enrichment processes.

b. Constructing an appropriate hypothesis about the intended contextual assumptions (in relevance-theoretic terms, IMPLICATED PREMISES).

c. Constructing an appropriate hypothesis about the intended contextual implications (in relevance-theoretic terms, IMPLICATED CONCLUSIONS).

These sub-tasks should not be thought of as sequentially ordered. The hearer does not FIRST decode the logical form of the sentence uttered, THEN construct an explicature and select an appropriate context, and THEN derive a range of implicated conclusions. Comprehension is an on-line process, and hypotheses about explicatures, implicated premises and implicated conclusions are developed in parallel against a background of expectations (or anticipatory hypotheses) which may be revised or elaborated as the utterance unfolds.[18] In particular, the hearer may bring to the comprehension process not only a general presumption of relevance, but more specific expectations about how the utterance will be relevant to him (what cognitive effects it is likely to achieve), and these may contribute, via backwards inference, to the identification of explicatures and implicated premises.[19] Thus, each sub-task in (10a-c) above involves a non-demonstrative inference process embedded within the overall process of constructing a hypothesis about the speaker's meaning.

To take just one illustration, consider the exchange in (11):

(11) a. *Peter:* Did John pay back the money he owed you?

b. *Mary:* No. He forgot to go to the bank.

Here is a schematic outline of how Peter might use the relevance-theoretic comprehension procedure to construct hypotheses about the explicatures and implicatures of Mary's utterance, "He forgot to go to the bank":

(12)

(a) Mary has said to Peter, "He_x forgot to go to the $BANK_1$ / $BANK_2$." [He_x = uninterpreted pronoun] [$BANK_1$ = financial institution] [$BANK_2$ = river bank]	*Embedding of the decoded (incomplete) logical form of Mary's utterance into a description of Mary's ostensive behaviour.*
(b) Mary's utterance will be optimally relevant to Peter.	*Expectation raised by recognition of Mary's ostensive behaviour and acceptance of the presumption of relevance it conveys.*
(c) Mary's utterance will achieve relevance by explaining why John has not repaid the money he owed her.	*Expectation raised by (b), together with the fact that such an explanation would be most relevant to Peter at this point.*
(d) Forgetting to go to the $BANK_1$ may make one unable to repay the money one owes.	*First assumption to occur to Peter which, together with other appropriate premises, might satisfy expectation (c). Accepted as an implicit premise of Mary's utterance.*
(e) John forgot to go to the $BANK_1$.	*First enrichment of the logical form of Mary's utterance to occur to Peter which might combine with (d) to lead to the satisfaction of (c). Accepted as an explicature of Mary's utterance.*
(f) John was unable to repay Mary the money he owes because he forgot to go to the $BANK_1$.	*Inferred from (d) and (e), satisfying (c) and accepted as an implicit conclusion of Mary's utterance.*
(g) John may repay Mary the money he owes when he next goes to the $BANK_1$.	*From (f) plus background knowledge. One of several possible weak implicatures of Mary's utterance which, together with (f), satisfy expectation (b).*

Peter assumes in (12b) that Mary's utterance, decoded as in (12a), is optimally relevant to him. Since what he wants to know at this point is why John did not repay the money he owed, he assumes in (c) that Mary's utterance will achieve relevance by answering this question. In the situation described, the logical form of the utterance provides easy access to the contextual assumption in (d) (that forgetting to go to the bank may prevent one from repaying money one owes). This could be used as an implicit premise in deriving the expected explanation of John's behaviour, provided that the utterance is interpreted on the explicit side (via disambiguation and reference resolution) as conveying the information in (e): that John forgot to go to the $BANK_1$. By combining the implicit premise in (d) and the explicit premise in (e), Peter arrives at the implicit conclusion in (f), from which further, weaker implicatures, including (g) and others, can be derived. The resulting overall interpretation satisfies Peter's expectations of relevance. On this account, explicatures and IMPLICATURES (i.e. implicit premises and conclusions) are arrived at by a process of mutual parallel adjustment, with hypotheses about both being considered in order of accessibility.[20]

This schematic outline of the comprehension process is considerably oversimplified.[21] In particular, it omits a range of lexical-pragmatic processes which contribute in important ways to the construction of explicatures. Consider the word *bank* in (11b). In interpreting this utterance, Peter would probably take Mary to be referring not just to a banking institution but to a specific type of banking institution: one that deals with private individuals, and in particular, with John. Unless the denotation of *bank* is narrowed in this way, the explicit content of Mary's utterance will not warrant the conclusion in (12f), which is needed to satisfy Peter's expectation of relevance. (It is hard to see how the fact that John had forgotten to go to the World Bank, say, or the European Investment Bank, might explain his failure to repay the money he owed.) By the

same token, in interpreting the phrase *go to the bank*, he would take Mary to be referring not merely to visiting the bank but to visiting it in order to get money, and, moreover, to get money in the regular way (legally, rather than, say, by robbing the bank). Unless the explicit content of the utterance is narrowed in this way, it will not warrant the conclusion in (12f), which is needed to satisfy Peter's expectation of relevance.

Some of these stereotypical narrowings have been described in the pragmatic literature as generalised conversational implicatures, and analysed as default interpretations, derivable via default rules.[22] Despite the richness and subtlety of much of the literature on generalised conversational implicature, relevance theory takes a different approach, for two main reasons. In the first place, as noted above, it treats lexical narrowing as a type of pragmatic enrichment process which contributes to explicatures rather than implicatures.[23] Like all enrichment processes, lexical narrowing is driven by the search for relevance, which involves the derivation of cognitive effects, and in particular of contextual implications. By definition, a contextual implication must follow logically from the explicatures of the utterance and the context. Sometimes, as in (11b), in order to yield an expected implication, the explicit content of the utterance must be enriched to a point where it warrants the expected conclusion. In any framework where implicated conclusions are seen as logically warranted by explicit content, there is thus good reason to treat lexical narrowings as falling on the explicit rather than the implicit side.[24]

In the second place, lexical narrowing is a much more flexible and context-dependent process than appeals to generalised implicature or default interpretations suggest. Barsalou (1987, 1992) surveys a wide range of experimental evidence which shows that even apparently stereotypical narrowings of terms such as *bird, animal, furniture, food,* etc. vary considerably across situations, individuals and times, and are strongly affected by discourse context and

considerations of relevance. In Barsalou's view, his results are best explained by the assumption that lexical items give access not to ready-made prototypes (assignable by default rules) but to a vast array of encyclopaedic information which varies in accessibility from occasion to occasion, with different subsets being selected ad hoc to determine the occasion-specific interpretation of a word. On this approach, *bank* in (11b) might be understood as conveying not the encoded concept $BANK_1$ but the related concept BANK*, with a more restricted encyclopedic entry and a narrower denotation, constructed ad hoc for this particular occasion.

In Barsalou's view, the construction of ad hoc concepts is affected by a variety of factors, including context, the accessibility of encyclopedic assumptions and considerations of relevance. However, he makes no concrete proposal about how these concepts might be derived, and in particular about how the construction process is triggered and when it stops. The relevance-theoretic comprehension procedure may be seen as a concrete hypothesis about how such a flexible, relevance-governed lexical interpretation process might go. The hearer treats the linguistically encoded word meaning (e.g. $BANK_1$ in (11b)) as no more than a clue to the speaker's meaning. Guided by his expectations of relevance, and using contextual assumptions made accessible by the encyclopedic entry of the linguistically encoded concept (e.g. that forgetting to go to the bank where one keeps one's money may make one unable to repay money one owes), he starts deriving cognitive effects. When he has enough effects to satisfy his expectations of relevance, he stops. The results would be as in (12) above, except that the contextual assumption in (d), the explicature in (e) and the implicatures in (f) and (g) would contain not the encoded concept $BANK_1$ but the ad hoc concept BANK*, with a narrower denotation, which would warrant the derivation of the cognitive effects required to satisfy the hearer's expectations of relevance.

The effect of such a flexible interpretation process may be a loosening rather than a narrowing of the encoded meaning (resulting in a broader rather than a narrower denotation). This is another way in which lexical pragmatic processes differ from default, or stereotypical, narrowing. Clear cases of loosening include the use of a prominent brand name (e.g. *Hoover, Xerox, Kleenex*) to denote a category which also contains items from less prominent brands; other good examples are approximations based on well-defined terms such as *square, painless* or *silent*, but the phenomenon is very widespread. Consider *bank* in (11b). Given current banking practice, the word may sometimes be loosely used to denote a category containing not only banking institutions but also the automatic cash dispensers found in supermarkets and stations. Indeed, in order to satisfy his expectations of relevance in (11b), Peter would probably have to take it in this way (i.e. to mean, roughly, 'bank-or-cash-dispenser'). (If John regularly gets his money from a cash dispenser, the claim that he forgot to go to the $BANK_{1,}$ might be strictly speaking false, and in any case would not adequately explain his failure to repay Mary.) Thus, *bank* in (11b) might be understood as expressing not the encoded concept $BANK_1$, but an ad hoc concept BANK**, with a broader denotation, which shares with $BANK_1$ the salient encyclopedic attribute of being a place one goes to in order to access money from one's account. The interpretation of a quite ordinary utterance such as (11b) might then involve both a loosening and a narrowing of the encoded meaning.

Loose uses of language present a problem for Grice's framework. Strictly speaking, faces are not square, rooms are generally not silent, and to describe them as such would violate his maxim of truthfulness ("Do not say what you believe to be false"). However, these departures from truthfulness do not fall into any of the categories of maxim-violation recognised by Grice (Grice 1989: 30). They are not covert violations, like lies, designed to deceive the hearer into believing what was said. They are not like jokes and fictions, which suspend the maxims entirely. Given

their intuitive similarities to metaphor and hyperbole, it might be tempting to analyse them, like tropes, as overt violations (floutings) of the maxim of truthfulness, designed to trigger the search for a related implicature (in this case, a hedged version of what was said). The problem is that these loose uses of language would not be generally perceived as violating the maxim of truthfulness at all. They do not have the striking quality that Grice associated with floutings, and which he saw as resulting in figurative or quasi-figurative interpretations. While we are all capable of realising on reflection that they are not strictly and literally true, these departures from truthfulness pass unattended and undetected in the normal flow of discourse. Grice's framework thus leaves them unexplained.[25]

Loose uses of language are not the only problem for Grice's maxim of truthfulness. There are questions about how the maxim itself is to be understood, and a series of difficulties with the analysis of tropes as overt violations of the maxim (for detailed discussion, see Wilson & Sperber 2002). Notice, too, that the intuitive similarities between loose talk, metaphor and hyperbole cannot be captured within this framework, since metaphor and hyperbole are seen as overt violations of the maxim of truthfulness, while loose uses of language are not. We have argued that the best solution is to abandon the maxim of truthfulness and treat whatever expectations of truthfulness arise in utterance interpretation as resulting not from an independent maxim, norm or convention of truthfulness, but as by-products of the more basic expectation of relevance. On this approach, loose talk, metaphor and hyperbole involve no violation of any maxim, but are merely alternative routes to achieving optimal relevance. Whether an utterance is literally, loosely or metaphorically understood will depend on the mutual adjustment of context, context and cognitive effects in the effort to satisfy the hearer's overall expectation of relevance.[26]

To illustrate this unified approach, consider the exchange in (13):

(13) a. *Peter*: What do you think of Martin's latest novel?

b. *Mary*: It puts me to sleep.

In Grice's framework, Mary's utterance in (13b) should have three distinct interpretations: as a literal assertion, a hyperbole or a metaphor.[27] Of these, Peter should test the literal interpretation first, and move to a figurative interpretation only if the literal interpretation blatantly violates the maxim of truthfulness. Yet there is now a lot of experimental evidence suggesting that literal interpretations do not have to be tested and rejected before figurative interpretations are considered;[28] indeed, in interpreting (13b), it would probably not even occur to Peter to wonder whether Mary literally fell asleep.

The relevance-theoretic analysis takes these points into account. In the first place, there is no suggestion that the literal meaning must be tested first. As with *bank* in (11b), the encoded conceptual address is treated merely as a point of access to an ordered array of encyclopedic information from which the hearer is expected to select in constructing a satisfactory overall interpretation. Whether this interpretation is literal or loose will depend on which types of information he selects. In processing (13b), Peter will be expecting to derive an answer to his question: that is, an evaluation of the book. In the circumstances, the first contextual assumption to occur to him is likely to be that a book which puts one to sleep is extremely boring and unengaging. Having used this assumption to derive an answer to his question, thus satisfying his expectations of relevance, he should stop. Just as in interpreting *bank* in (11b), it does not occur to him to wonder whether John gets his money from a bank or a cash dispenser, so in interpreting (13b), it should not occur to him to wonder whether the book literally puts Mary to sleep, almost puts her to sleep or merely bores her greatly. Just as the mutual adjustment process in (13) yields an explicature containing the ad hoc concept BANK**, which has undergone simultaneous

narrowing and loosening, so the mutual adjustment process for (13b) should yield an explicature containing the ad hoc concept PUT TO SLEEP*, which denotes not only literal cases of putting to sleep, but other cases that share with it the encyclopedic property of being extremely boring and unengaging. Only if such a loose interpretation fails to satisfy his expectations of relevance would Peter be justified in spending the effort required to explore further contextual assumptions, and moving towards a more literal interpretation.[29]

Typically, the explicit content of loose uses in general, and of metaphors in particular, exhibits a certain degree of indeterminacy. Compare, for instance, the results of using the word *square* literally in a geometric statement to convey the concept SQUARE, using it loosely in the phrase *a square face* to convey the concept SQUARE*, and using it metaphorically in the phrase *a square mind* to convey the concept SQUARE**. In relevance theory, this relative indeterminacy of explicatures is linked to the relative strength of implicatures.

A proposition may be more or less strongly implicated by an utterance. It is STRONGLY IMPLICATED (or is a STRONG IMPLICATURE) if its recovery is essential in order to arrive at an interpretation that satisfies the expectations of relevance raised by the utterance itself. It is WEAKLY IMPLICATED if its recovery helps with the construction of an interpretation that is relevant in the expected way, but is not itself essential because the utterance suggests a range of similar possible implicatures, any one of which would do (Sperber & Wilson 1986a: §1.10-12, §4.6). For instance, (11b), "He forgot to go to the bank", strongly implicates (12f), *John was unable to repay Mary the money he owes her because he forgot to go to the BANK$_1$*, since without this implication,[30] (11b) is not a relevant reply to (11a), "Did John pay back the money he owed you?" (11b) also encourages the audience to derive a further implicature along the lines of (12g), *John may repay Mary the money he owes her when he next goes to the BANK$_1$*, but here the audience must take some responsibility for coming to this conclusion rather than, say, the

conclusion that John WILL repay Mary the money he owes her when he next goes to the $BANK_1$, or some other similar conclusion.

Typically, loose uses, and metaphorical uses in particular, convey an array of weak implicatures. Thus, the utterance *John has a square mind* weakly implicates that he is somewhat rigid in his thinking, that he does not easily change his mind, that he is a man of principle, and so on. None of these implicatures is individually required for the utterance to make sense, but, on the other hand, without some such implicatures, it will make no sense at all. If the word *square* is understood as conveying the concept SQUARE**, which combines with contextual information to yield these implications, then the concept SQUARE** itself will exhibit some indeterminacy or fuzziness, and the utterance as a whole will exhibit a corresponding weakness of explicature. Loose uses and metaphors typically exhibit such fuzziness, for which relevance theory provides an original account.

The distinction between strong and weak implicatures sheds some light on the variety of ways in which an utterance can achieve relevance. Some utterances (technical instructions, for instance) achieve relevance by conveying a few strong implicatures. Other utterances achieve relevance by weakly suggesting a wide array of possible implications, each of which is a weak implicature of the utterance. This is typical of poetic uses of language, and has been discussed in relevance theory under the heading of POETIC EFFECT (Sperber & Wilson 1986a: §4.6-9; Pilkington 2000).

In Grice's framework (and indeed in all rhetorical and pragmatic discussions of irony as a figure of speech before Sperber & Wilson 1981) the treatment of verbal irony parallels the treatments of metaphor and hyperbole. For Grice, irony is an overt violation of the maxim of truthfulness, and differs from metaphor and hyperbole only in the kind of implicature it conveys (metaphor implicates a simile based on what was said, hyperbole implicates a weakening of what

was said, and irony implicates the opposite of what was said). Relevance theorists have argued against not only the Gricean analysis of irony but the more general assumption that metaphor, hyperbole and irony should be given parallel treatments.

Grice's analysis of irony as an overt violation of the maxim of truthfulness is a variant of the classical rhetorical view of irony as literally saying one thing and figuratively meaning the opposite. There are well-known arguments against this view. It is descriptively inadequate because ironical understatements, ironical quotations and ironical allusions cannot be analysed as communicating the opposite of what is literally said. It is theoretically inadequate because saying the opposite of what one means is patently irrational; and on this approach it is hard to explain why verbal irony is universal and appears to arise spontaneously, without being taught or learned (Sperber & Wilson 1981, 1998b; Wilson & Sperber 1992).

Moreover, given the relevance-theoretic analysis of metaphor and hyperbole as varieties of loose use, the parallelism between metaphor, hyperbole and irony cannot be maintained. While it is easy to see how a speaker aiming at optimal relevance might convey her meaning more economically by speaking loosely rather than using a cumbersome literal paraphrase, it is hard to see how a rational speaker could hope to convey her meaning more economically by choosing a word whose encoded meaning is the opposite of the one she intends to convey (or how a hearer using the relevance-theoretic comprehension procedure could understand her if she did). Some alternative explanation of irony must be found.

According to the explanation proposed by relevance theory, verbal irony involves no special machinery or procedures not already needed to account for a basic use of language, INTERPRETIVE USE, and a specific form of interpretive use, ECHOIC USE. An utterance may be interpretively used to (meta)represent another utterance or thought that it resembles in content. The best-known type of interpretive use is reported speech or thought. An utterance is echoic when it achieves most of

its relevance not by expressing the speaker's own views, nor by reporting someone else's utterances or thoughts, but by expressing the speaker's attitude to views she tacitly attributes to someone else.[31] To illustrate, suppose that Peter and Mary are leaving a party, and one of the following exchanges occurs:

(14) *Peter*: That was a fantastic party.

(15) *Mary*: a. [happily] Fantastic.

b. [puzzled] Fantastic?

c. [scornfully] Fantastic!

In (15a), Mary echoes Peter's utterance in order to indicate that she agrees with it; in (15b), she indicates that she is wondering about it; and in (15c) she indicates that she disagrees with it. The resulting interpretations might be as in (16):

(16) a. She believes I was right to say/think that the party was fantastic.

b. She is wondering whether I was right to say/think that the party was fantastic.

c. She believes I was wrong to say/think that the party was fantastic.

Here, the basic proposition expressed by the utterances in (15) (that *the party was fantastic*) is embedded under an appropriate higher-order speech-act or propositional-attitude description indicating, on the one hand, that the basic proposition is being used to interpret views Mary attributes to someone else, and, on the other, Mary's attitude to these views. In order to understand Mary's meaning, Peter has to recognise not only the basic proposition expressed but also the fact that it is being attributively used, and Mary's attitude to the attributed views.

The attitudes conveyed by use of an echoic utterance may be very rich and varied. The speaker may indicate that she endorses or dissociates herself from the thought or utterance she is

echoing: that she is puzzled, angry, amused, intrigued, sceptical, and so on, or any combination of these. On the relevance-theoretic account, verbal irony involves the expression of a tacitly dissociative attitude – wry, sceptical, bitter or mocking – to an attributed utterance or thought. Consider Mary's utterance in (15c) above. This is clearly both ironical and echoic. Relevance theory claims that it is ironical BECAUSE it is echoic: verbal irony consists in echoing a tacitly attributed thought or utterance with a tacitly dissociative attitude.[32]

This approach sheds light on many cases of irony not dealt with on the classical or Gricean accounts. Consider Mary's utterance "He forgot to go to the bank" in (11b) above. There are situations where this might well be ironical, even though it is neither blatantly false nor used to convey the opposite of what was said. Suppose Peter and Mary both know that John has repeatedly failed to repay Mary, with a series of pitifully inadequate excuses. Then (11b) may be seen as an ironical echo in which Mary tacitly dissociates herself from the latest excuse in the series. Thus, all that is needed to make (11b) ironical is a scenario in which it can be understood as a mocking echo of an attributed utterance or thought.[33]

One implication of this analysis is that irony involves a higher order of metarepresentational ability than metaphor. On the relevance-theoretic account, as illustrated in (16) above, the interpretation of echoic utterances in general involves the ability to recognise that the speaker is thinking, not directly about a state of affairs in the world, but about another thought or utterance that she attributes to someone else. This implication of our account is confirmed by experimental evidence from the literature on autism, child development and right hemisphere damage, which shows that the comprehension of irony correlates with second-order metarepresentational abilities, while the comprehension of metaphor requires only first-order abilities.[34] This fits straightforwardly with the relevance-theoretic account of irony, but is unexplained on the classical or Gricean accounts.[35]

Another area in which metarepresentational abilities play an important role is the interpretation of illocutionary acts. Consider the exchange in (17):

(17) a. *Peter:* Will you pay back the money by Tuesday?

b. *Mary:* I will pay it back by then.

Both (17a) and (17b) express the proposition that *Mary will pay back the money by Tuesday*. In the interrogative (17a), this proposition is expressed but not communicated (in the sense that Peter does not put it forward as true, or probably true)[36]: in relevance-theoretic terms, it is not an explicature of Peter's utterance. Yet intuitively, (17a) is no less explicit an act of communication than (17b). Relevance theory claims that what is explicitly communicated by (17a) is the higher-order speech-act description in (18):

(18) Peter is asking Mary whether she will pay back the money by Tuesday;

Like all explicatures, (18) is recovered by a mixture of decoding and inference based on a variety of linguistic and non-linguistic clues (e.g. word order, mood indicators, tone of voice, facial expression): in relevance-theoretic terms, it is a HIGHER-LEVEL EXPLICATURE of (17a).[37] In (17b), by contrast, the explicatures might include both (19a), the BASIC EXPLICATURE, and higher-level explicatures such as (19b) and (19c):

(19) a. Mary will pay back the money by Tuesday.

b. Mary is promising to pay back the money by Tuesday.

c. Mary believes she will pay back the money by Tuesday.

Thus, an utterance may convey several explicatures, each of which may contribute to relevance and warrant the derivation of implicatures.[38]

On this approach, verbal irony has more in common with illocutionary and attitudinal utterances than it does with metaphor or hyperbole. The recognition of irony, like the recognition of illocutionary acts and expressions of attitude, involves a higher order of metarepresentational ability than the recognition of the basic proposition expressed by an utterance, whether literal, loose or metaphorical. More generally, on both Gricean and relevance-theoretic accounts, the interpretation of every utterance involves a high degree of metarepresentational capacity, since comprehension rests on the ability to attribute both informative and communicative intentions. This raises the question of how pragmatic abilities are acquired, and how they fit into the overall architecture of the mind.

5. Relevance theory and mental architecture

Grice's analysis of overt communication as involving the expression and recognition of intentions treats comprehension as a variety of MIND-READING, or THEORY OF MIND (the ability to attribute mental states to others in order to explain and predict their behaviour).[39] The link between mind-reading and communication is confirmed by a wealth of developmental and neuropsychological evidence.[40] However, mind-reading itself has been analysed in rather different ways. Philosophers often describe it as an exercise in reflective reasoning (a central thought process, in the terms of Fodor 1983), and many of Grice's remarks about pragmatics are consistent with this. Thus, his rational reconstruction of how conversational implicatures are derived is a straightforward exercise in "belief-desire" psychology, involving the application of general-purpose reasoning mechanisms to premises based on explicit hypotheses about the relations between mental states and behaviour:

> He said that P; he could not have done this unless he thought that Q; he knows (and knows that I know that he knows) that I will realise that it is necessary to suppose that Q; he has done nothing to stop me thinking that Q; so he intends me to think, or is at least willing for me to think, that Q. (Grice, 1989: 30–31)

In our own early work, we also treated pragmatic interpretation as a central, inferential process (as opposed to part of a peripheral language module), albeit a spontaneous, intuitive rather than a conscious, reflective one (Sperber & Wilson 1986a: chapter 2; Wilson & Sperber 1986). More recently, there has been a tendency in the cognitive sciences to move away from Fodor's sharp distinction between modular input processes and relatively undifferentiated central processes and towards an increasingly modular view of the mind.[41] In this section, we will consider how the relevance-theoretic comprehension procedure might fit with more modular accounts of inference, and in particular of mind-reading.[42]

One advantage of a dedicated inferential mechanism or module is that it can take advantage of regularities in its own particular domain, and contain special-purpose inferential procedures which are justified by these regularities, but only in this domain. Thus, in modular accounts of mind-reading, standard "belief-desire" psychology may be replaced by special-purpose inferential procedures ("fast and frugal heuristics", in the terms of Gigerenzer et al. 1999) attuned to the properties of this particular domain. Examples discussed in the literature on mind-reading include an Eye Direction Detector which attributes perceptual and attentional states on the basis of direction of gaze, and an Intentionality Detector which interprets self-propelled motion in terms of goals and desires (Leslie 1994; Premack & Premack 1994; Baron-Cohen1995). In mechanisms of this "fast and frugal" type, regularities in the relations between

mental states and behaviour are not registered as explicit premises in an inference process, but function merely as tacit underpinnings for the working of the device.

Most approaches to mind-reading, whether modular or non-modular, have tended to take for granted that there is no need for special-purpose inferential comprehension procedures, because the mental-state attributions required for comprehension will be automatically generated by more general mind-reading mechanisms which apply across the whole domain.[43] We believe that there are serious problems with the view that speakers' meanings can be inferred from utterances by the same procedures used to infer intentions from actions. In the first place, the range of actions an agent can reasonably intend to perform in a given situation is in practice quite limited. Regular intention attribution is greatly facilitated by the relatively narrow range of actions available to an agent at a time. By contrast, as noted above (§3), the range of meanings a speaker can reasonably intend to convey in a given situation is virtually unlimited. It is simply not clear how the standard procedures for intention attribution could yield attributions of speakers' meanings except in easy and trivial cases (for further discussion, see Sperber 2000; Sperber & Wilson 2002).

In the second place, as noted above (§4), inferential comprehension typically involves several layers of metarepresentation, while in regular mind-reading a single level is generally enough. This discrepancy between the metarepresentational capacities required for inferential comprehension and regular mind-reading is particularly apparent in child development. It is hard to believe that two-year-old children, who fail on regular first-order false belief tasks, can recognise and understand the peculiar multi-levelled representations involved in verbal comprehension, using nothing more than a general ability to attribute intentions to agents in order to explain their behaviour. For these reasons, it is worth exploring the possibility that, within the overall mind-reading module, there has evolved a specialised sub-module dedicated to

comprehension, with its own proprietary concepts and mechanisms (Sperber 1996, 2000, 2002; Origgi & Sperber 2000; Wilson 2000; Sperber & Wilson 2002).

If we are right, the Communicative Principle of Relevance in (7) above ("Every ostensive stimulus conveys a presumption of its own optimal relevance") describes a regularity specific to the communicative domain. Only acts of ostensive communication create legitimate presumptions of optimal relevance, and this might form the basis for a special-purpose inferential comprehension device. On this modular account, the relevance-theoretic comprehension procedure in (9) above ("Follow a path of least effort in computing cognitive effects: test interpretive hypotheses in order of accessibility; stop when your expectations of relevance are satisfied") could be seen as a "fast and frugal heuristic" which automatically computes a hypothesis about the speaker's meaning on the basis of the linguistic and other evidence provided.

The complexity of the inferences required on the Gricean account of communication has sometimes been seen as an argument against the whole inferential approach. We are suggesting an alternative view on which, just as children do not have to learn their language but come with a substantial innate endowment, so they do not have to learn what ostensive-inferential communication is, but come with a substantial innate endowment. This approach allows for varying degrees of sophistication in the expectations of relevance with which an utterance is approached. In the terms of Sperber (1994), a child with limited metarepresentational capacity might start out as a Naively Optimistic interpreter, who accepts the first interpretation he finds relevant enough regardless of whether it is one the speaker could plausibly have intended. A Cautious Optimist, with enough metarepresentational capacity to pass first-order false belief tasks, might be capable of dealing with mismatches of this type, but unable to deal with deliberate deception (Sperber 1994; Bezuidenhout & Sroda 1998; Wilson 2000; Happé & Loth

2002). A Sophisticated Understander has the metarepresentational capacity to deal simultaneously with mismatches and deception. In the relevance-theoretic framework, normal adults are seen as Sophisticated Understanders, and this is an important difference from the standard Gricean approach (for references and discussion, see footnotes 9 and 19).

6. Conclusion: an experimentally testable cognitive theory

Relevance theory is a cognitive psychological theory. In particular, it treats utterance interpretation as a cognitive process. Like other psychological theories, it has testable consequences: it can suggest experimental research, and is open to confirmation, disconfirmation or fine-tuning in the light of experimental evidence. Of course, as with other theories of comparable scope, its most general tenets can be tested only indirectly, by evaluating some of their consequences. Thus, the Cognitive Principle of Relevance (the claim that human cognition tends to be geared to the maximisation of relevance) suggests testable predictions only when combined with descriptions of particular cognitive mechanisms (for perception, categorisation, memory, or inference, for example). Given a description of such a mechanism, it may be possible to test the relevance-theoretic claim that this mechanism contributes to a greater allocation of cognitive resources to potentially relevant inputs, by comparing it with some alternative hypothesis, or at least the null hypothesis.

The Communicative Principle of Relevance (the claim that every ostensive stimulus conveys a presumption of its own optimal relevance) is a law-like generalisation which follows from the Cognitive Principle of Relevance, together with a broadly Gricean view of communication as a process of inferential intention-attribution. The Communicative Principle of Relevance could be falsified by finding genuine communicative acts which did not convey a presumption of their own optimal relevance (but conveyed instead, say, a presumption of literal truthfulness, or

maximal informativeness, or no such presumption at all). When combined with descriptions of specific types and properties of communicative acts (and in particular of utterances), the Communicative Principle yields precise predictions, some of which have been experimentally tested.

Throughout this survey, we have tried to point out cases where the predictions of relevance theory differ from those more or less clearly suggested by alternative frameworks (e.g. on the interpretation of ostensive silences, the order of accessibility of literal and metaphorical interpretations, the contribution of pragmatic principles to explicit communication, the nature of lexical-pragmatic processes, the parallelism between metaphor and irony), and we have drawn attention to many cases where the relevance-theoretic analyses have been experimentally tested and their predictions confirmed. Here we will give two further illustrations of how the basic notion of optimal relevance, characterised in terms of effort and effect, yields testable predictions.

As noted above (§2), relevance theory does not provide an absolute measure of mental effort or cognitive effect, and it does not assume that such a measure is available to the spontaneous workings of the mind. What it does assume is that the actual or expected relevance of two inputs can quite often be compared. These possibilities of comparison help individuals to allocate their cognitive resources, and communicators to predict and influence the cognitive processes of others. They also make it possible for researchers to manipulate the effect and effort factors in experimental situations.

Thus, consider a conditional statement such as “If a card has a 6 on the front, it has an E on the back.” In the Wason selection task (Wason 1966), the most famous experimental paradigm in the psychology of reasoning, participants are presented with four cards showing (say) a 6, a 4, an E and an A, and asked which of these cards should be turned over in order to reveal the hidden

letter or number and check whether the conditional statement is true or false. The correct response is to select the 6 and A cards. By 1995, literally thousands of experiments with such materials had failed to elicit a majority of correct responses. Most people choose the 6 alone, or the 6 and the E. In “Relevance theory explains the selection task” (1995), Sperber, Cara and Girotto argued that participants derive testable implications from the conditional statement in order of accessibility, stop when their expectations of relevance are satisfied, and select cards on the basis of this interpretation. Using this idea, Sperber et al. were able, by varying the content and context of the conditional statement, to manipulate the effort and effect factors so as to elicit correct or incorrect selections at will.

Typically, a conditional statement *If P then Q* achieves relevance by making it possible to derive the consequent *Q* in cases where the antecedent *P* is satisfied. With the conditional “If a card has a 6 on the front, it has an E on the back”, this leads to selection of the card with a 6. Another common way for a conditional statement to achieve relevance is by creating an expectation that the antecedent *P* and the consequent *Q* will both be true. In the present case, this leads to selection of the 6 and E cards. Of course, a conditional statement also implies that its antecedent and the negation of its consequent will not be true together. If participants chose cards on this basis, they would correctly select the 6 and A cards. However, in most contexts this implication is relatively costly to derive, yields no further effects, and would not be derived by a hearer looking for optimal relevance. What Sperber et al. did was to manipulate the effort and effect factors, either separately or together, in such a way that this implication was easier and/or more rewarding to derive, and the correct cards were therefore increasingly likely to be chosen. The most successful condition was one in which the statement “If a card has a 6 on the front, it has an E on the back” was seen as coming from an engineer who had just repaired a machine which was supposed to print cards with this specification, but which had malfunctioned and

printed cards with a 6 on the front and an A on the back. Here, the statement achieved relevance by implying that there would be no more cards with a 6 and an A rather than an E, and a majority of participants made the correct selection. This and other experiments with the selection task (see also Girotto, Kemmelmeir, Sperber & van der Henst 2001; Sperber & Girotto forthcoming) showed that performance on this task was determined not by domain-general or domain-specific reasoning mechanisms (as had been argued by most reseachers) but by pragmatic factors affecting the interpretation of conditional statements. It also confirmed that the interpretation of conditionals is governed by the twin factors of effort and effect, which can act either separately or in combination.[44]

Here is a second example of how the interaction of effort and effect can be experimentally investigated, this time in the production rather than the interpretation of utterances. Suppose a stranger comes up and asks me the time. I look at my watch and see that it is 11:58 exactly. How should I reply? A speaker observing Grice's maxims (and in particular the maxim of truthfulness), addressing an audience who expects her to observe these maxims, should say, "11:58". In this situation, a speaker who rounds up the time to 12:00 (thus speaking loosely and violating the maxim of truthfulness) would create the mistaken assumption that she meant to convey that it was (exactly) 12:00. By contrast, a speaker aiming at optimal relevance would have every reason to round up the time (thus reducing her hearer's processing effort) unless, in her view, some cognitive effect would be lost by speaking loosely. It should therefore be possible, on the relevance-theoretic approach, to elicit stricter or looser answers by manipulating the scenario in which the question is asked, so that the stricter answer does or does not yield relevant implications. This prediction has been experimentally tested, and the relevance-theoretic analysis confirmed: strangers in public places asked for the time tend either to round up or to give answers

that are accurate to the minute, depending on subtle clues as to what might make it relevant for the person making the request to know the time (van der Henst, Carles & Sperber forthcoming).

Currently, the main obstacle to experimental comparisons of relevance theory with other pragmatic theories is that the testable consequences of these other theories have often not been made explicit. Much pragmatic research has been carried out in a philosophical or linguistic tradition in which the goal of achieving theoretical generality, combined with a tendency to rely on intuitions, has created a certain reluctance to get down to the messy business of experimentation. Relevance theorists have been trying to combine theoretical generality with all the possibilities of testing provided by the careful use of linguistic intuitions, observational data, and the experimental methods of cognitive psychology. We see this as an important direction for future research.

References

Asher, Nicholas & Lascarides, Alex. 2001. The semantics and pragmatics of metaphor. In P. Bouillon & F. Busa (eds) *The Language of Word Meaning*: 262-289. Cambridge University Press, Cambridge.

Bach, Kent. 1994a. Conversational impliciture. *Mind and Language* 9: 124-62.

Bach, Kent. 1994b. Semantic slack: What is said and more. In S. Tsohatzidis (ed.) *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*: 267-91. Routledge, London.

Bach, Kent. 1997. The semantics–pragmatics distinction: What it is and why it matters. *Linguistische Berichte* 8 (Special Issue on Pragmatics): 33-50. Reprinted in K. Turner (ed). 1999. *The Semantics–Pragmatics Interface from Different Points of View*. Elsevier Science, Oxford.

Barkow, John, Cosmides, Leda & Tooby, John. 1995. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press, Oxford.

Baron-Cohen, Simon. 1995. *Mindblindness: An Essay on Autism and Theory of Mind.* MIT Press, Cambridge MA.

Barsalou, Lawrence. 1987. The instability of graded structure: implications for the nature of concepts. In U. Neisser (ed.) *Concepts and Conceptual Development*: *Ecological and Intellectual Factors in Categorization*: 101-140. Cambridge University Press, Cambridge.

Barsalou, Lawrence. 1992. Frames, concepts, and conceptual fields. In E. Kittay & A. Lehrer (eds.) *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*: 21-74. Lawrence Erlbaum, Hillsdale NJ.

Bezuidenhout, Anne. 1997. Pragmatics, semantic underdetermination and the referential–attributive distinction. *Mind* 106: 375-409.

Bezuidenhout, Anne & Sroda, Mary Sue. 1998. Children's use of contextual cues to resolve referential ambiguity: An application of relevance theory. *Pragmatics and Cognition* 6: 265-299

Bird, Graham. 1994. Relevance theory and speech acts. In S. Tsohatzidis (ed.) *Foundations of Speech-Act Theory: Philosophical and Linguistic Perspectives*: 292-311. Routledge, London.

Blakemore, Diane. 1987. *Semantic Constraints on Relevance.* Blackwell, Oxford.

Blakemore, Diane. 1991. Performatives and parentheticals. *Proceedings of the Aristotelian Society* 91: 197-214.

Blakemore, Diane. 1994. Echo questions: A pragmatic account. *Lingua* 4: 197-211.

Blakemore, Diane. forthcoming. *Meaning and Relevance: The Semantics and Pragmatics of Discourse Connectives.* Cambridge University Press, Cambridge.

Blass, Regina. 1990. *Relevance Relations in Discourse: A Study with Special Reference to Sissala.* Cambridge University Press, Cambridge.

Bloom, Paul. 2000. *How Children Learn the Meanings of Words.* MIT Press, Cambridge MA.

Bloom, Paul. 2002. Mindreading, communication and the learning of names for things. *Mind & Language* 17: 37-54. Special Issue on Pragmatics and Cognitive Science.

Blutner, Reinhart. 1998. Lexical pragmatics. *Journal of Semantics* 15: 115-62.

Blutner, Reinhart. 2002. Lexical semantics and pragmatics. To appear in *Linguistische Berichte* 10.

Breheny, Richard. 2002. The current state of (radical) pragmatics in the cognitive sciences. *Mind and Language* 17: 169-87. Special Issue on Pragmatics and Cognitive Science.

Cappelen, Herman & Lepore, Ernie. 1997. Varieties of quotation. *Mind* 106: 429-450.

Carnap, Rudolf. 1950. *Logical Foundations of Probability.* Routledge & Kegan Paul, London.

Carruthers, Peter & Smith, Peter (eds.). 1996. *Theories of Theories of Mind.* Cambridge University Press, Cambridge.

Carston, Robyn. 1988. Implicature, explicature and truth-theoretic semantics. In R. Kempson (ed.) *Mental Representation: The Interface between Language and Reality.* Cambridge University Press, Cambridge: 155-81. Reprinted in Steven Davis (ed.) 1991: 33-51; Asa Kasher (ed.) 1998, vol. IV: 464-479.

Carston, Robyn. 1995. Quantity maxims and generalised implicatures. *Lingua* 96: 213-24.

Carston, Robyn. 1996. Metalinguistic negation and echoic use. *Journal of Pragmatics* 25: 309-330.

Carston, Robyn. 1997. Enrichment and loosening: complementary processes in deriving the proposition expressed? *Linguistische Berichte* 8 (Special Issue on Pragmatics): 103-127.

Carston, Robyn. 1998a. Informativeness, relevance and scalar implicature. In Robyn Carston & Seiji Uchida (eds.) 1998: 179-236.

Carston, Robyn. 1998b. *Pragmatics and the Explicit–Implicit Distinction.* University College London Ph.D. thesis.

Carston, Robyn. 1999. The semantics/pragmatics distinction: A view from relevance theory. In K. Turner (ed.) *The Semantics/Pragmatics Interface from Different Points of View* (CRiSPI 1): 85-125. Elsevier Science, Oxford.

Carston, Robyn. 2000. Explicature and semantics. *UCL Working Papers in Linguistics* 12: 1-44. To appear in S. Davis & B. Gillon (eds.) *Semantics: A Reader.* Oxford University Press, Oxford.

Carston, Robyn. 2002. Linguistic meaning, communicated meaning and cognitive pragmatics. *Mind & Language* 17: 127-148. Special Issue on Pragmatics and Cognitive Science.

Carston, Robyn. forthcoming. *Thoughts and Utterances: The Pragmatics of Explicit Communication.* Blackwell, Oxford.

Carston, Robyn & Uchida, Seiji (eds). 1998. *Relevance Theory: Applications and Implications.* John Benjamins, Amsterdam.

Clark, Herbert, and Gerrig, Richard. 1984. On the pretense theory of irony. *Journal of Experimental Psychology: General* 113: 121-6

Copestake, Anne & Briscoe, Ted. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12: 15-67.

Curcó, Carmen. 1998. Indirect echoes and verbal humour. In Villy Rouchota & Andreas Jucker (eds.) 1998: 305-325. John Benjamins, Amsterdam.

Davis, Steven (ed.). 1991. *Pragmatics: A Reader*. Oxford University Press, Oxford.

Davies, Martin, & Stone, Tony (eds.). 1995a. *Mental Simulation: Philosophical and Psychological Essays.* Blackwell, Oxford.

Davies, Martin, & Stone, Tony (eds.). 1995b. *Folk Psychology.* Blackwell, Oxford.

Fodor, Jerry. 1983. *The Modularity of Mind*. MIT Press, Cambridge MA.

Fodor, Jerry. 2000. *The Mind Doesn't Work That Way.* MIT Press, Cambridge MA.

Fodor, Jerry. 2001. Language, thought and compositionality. *Mind and Language* 16: 1-15.

Franks, Bradley & Braisby, Nicholas. 1990. Sense generation or how to make a mental lexicon flexible. *Proceedings of the 12th Annual Conference of the Cognitive Science Society.* Cambridge MA: MIT, July 1990.

Garnham, Alan, & Perner, Josef. 1990. Does manifestness solve problems of mutuality? *Behavioral and Brain Sciences* 13: 178-9.

Gibbs, Ray. 1994. *The Poetics of Mind: Figurative Thought, Language and Understanding*. Cambridge University Press, Cambridge.

Gibbs, Ray & Moise, Jessica. 1997. Pragmatics in understanding what is said. *Cognition* 62: 51-74.

Gibbs, Ray & O'Brien, J. 1991. Psychological aspects of irony understanding. *Journal of Pragmatics* 16: 523-30.

Gigerenzer, Gerd, Todd, Peter & the ABC Research Group. 1999. *Simple Heuristics that Make us Smart*. Oxford University Press, Oxford.

Girotto, Vittorio, Kemmelmeir, Markus, Sperber, Dan, & van der Henst, Jean-Baptiste. 2001. Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition*, 81: 69-76.

Glucksberg, Sam. 2001. *Understanding Figurative Language*. Oxford University Press, Oxford.

Green, Mitchell. 1995. Quantity, volubility, and some varieties of discourse. *Linguistics & Philosophy* 18: 83-112.

Grice, H. Paul. 1961. The causal theory of perception. *Proceedings of the Aristotelian Society, Supplementary Volume* 35: 121-152. Partially reprinted in H. Paul Grice 1989: 224-247.

Grice, H. Paul. 1989. *Studies in the Way of Words*. Harvard University Press, Cambridge MA.

Gutt, Ernst-August. 1991. *Translation and Relevance: Cognition and Context*. Blackwell, Oxford. (Second edition 2000. St Jerome Publishing, Manchester.)

Happé, Francesca. 1993. Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition* 48: 101-19

Happé, Francesca & Loth, Eva. 2002. 'Theory of mind' and tracking speakers' intentions. *Mind & Language* 17: 24-36. Special Issue on Pragmatics and Cognitive Science.

Harnish, Robert M. 1994. Mood, meaning and speech acts. In S. Tsohatzidis (ed.) *Foundations of Speech-Act Theory*: *Philosophical and Linguistic Perspectives*: 407-459. Routledge, London.

Hirschfeld, Laurence & Gelman, Susan (eds). 1994. *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge University Press, Cambridge.

Hobbs, Jerry. 1985. On the coherence and structure of discourse. CSLI Report 85-37. CSLI, Menlo Park CA.

Hobbs, Jerry, Stickel, Mark, Appelt, Douglas & Martin, Paul 1993. Interpretation as abduction. *Artificial Intelligence* 63: 69-142.

Horn, Laurence. 1984. Towards a new taxonomy for pragmatic inference: Q- and R-based implicature. In D. Schiffrin (ed.) *Meaning, Form, and Use in Context*: 11-42. Georgetown University Press, Washington DC.

Horn, Laurence. 1989. *A Natural History of Negation*. University of Chicago Press, Chicago IL.

Horn, Laurence. 1992. The said and the unsaid. *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory*: 163-202. Ohio State University Linguistics Department, Columbus OH.

Ifantidou, Elly. 2001. *Evidentials and Relevance*. John Benjamins, Amsterdam.

Jorgensen, Julia, Miller, George & Sperber, Dan. 1984. Test of the mention theory of irony. *Journal of Experimental Psychology: General* 113: 112-20

Kasher, Asa. 1976. Conversational maxims and rationality. In A. Kasher (ed.) *Language in Focus: Foundations, Methods and Systems*: 197-211. Reidel, Dordrecht. Reprinted in Asa Kasher (ed.) 1998, vol. IV.

Kasher, Asa. 1984. Pragmatics and the modularity of mind. *Journal of Pragmatics* 8: 539-557. Revised version reprinted in: Steven Davis (ed.) 1991: 567-582.

Kasher, Asa (ed.). 1998. *Pragmatics: Critical Concepts*, vols I-V. Routledge, London.

Kempson, Ruth. 1986. Ambiguity and the semantics–pragmatics distinction. In C. Travis (ed.) *Meaning and Interpretation*: 77-104. Blackwell, Oxford.

Kempson, Ruth. 1996. Semantics, pragmatics and deduction. In S. Lappin (ed.) *Handbook of Contemporary Semantic Theory*: 561-598. Blackwell, Oxford.

Kempson, Ruth & Cormack, Annabel. 1982. Ambiguity and quantification. *Linguistics and Philosophy* 4: 259-309.

Kreuz, Roger & Glucksberg, Sam. 1989. How to be sarcastic: The echoic reminder theory of irony. *Journal of Experimental Psychology: General* 118: 374-86.

Kumon-Nakamura, Sachi, Glucksberg, Sam & Brown, Mary. 1995. How about another piece of pie: the allusional pretense theory of discourse irony. *Journal of Experimental Psychology: General* 124: 3—21.

Lakoff, George. 1987. *Women, Fire and Dangerous Things.* University of Chicago Press, Chicago IL.

Langdon, Robyn, Davies, Martin & Coltheart, Max. 2002. Understanding minds and understanding communicated meanings in schizophrenia. *Mind & Language* 17: 68-104. Special Issue on Pragmatics and Cognitive Science.

Lascarides, Alex & Asher, Nicholas. 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics & Philosophy* 16: 437-93.

Lascarides, Alex & Copestake, Anne. 1998. Pragmatics and word meaning. *Journal of Linguistics* 34: 387-414.

Lascarides, Alex, Copestake, Anne, & Briscoe, Ted. 1996. Ambiguity and coherence. *Journal of Semantics* 13: 41-65.

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75: 522-551.

Leslie, Alan. 1991. The theory of mind impairment in autism: Evidence for a modular mechanism of development? In Andrew Whiten (ed.) 1991: 63-78.

Leslie, Alan. 1994. ToMM, ToBy, and Agency: Core architecture and domain specificity. In Laurence Hirschfeld & Susan Gelman (eds) 1994: 119-148.

Levinson, Stephen. 1987. Minimization and conversational inference. In J. Verschueren & M. Bertuccelli-Papi (eds) *The Pragmatic Perspective*: 61-129. John Benjamins, Amsterdam.

Levinson, Stephen. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. MIT Press, Cambridge MA.

Malle, Bertram, Moses, Louis & Baldwin, Dare (eds). 2001. *Intentions and Intentionality: Foundations of Social Cognition.* MIT Press, Cambridge MA.

Martin, Robert. 1992. Irony and universe of belief. *Lingua* 87: 77-90.

Matsui, Tomoko. 1998. Pragmatic criteria for reference assignment: A relevance-theoretic account of the acceptability of bridging. *Pragmatics and Cognition* 6: 47-97.

Matsui, Tomoko. 2000. *Bridging and Relevance*. John Benjamins, Amsterdam.

Matsumoto, Y. 1995. The conversational condition on Horn scales. *Linguistics and Philosophy* 18: 21-60.

Merin, Arthur. 1997. Information, relevance and social decisionmaking. In L. Moss, J. Ginzburg & M. De Rijke (eds.) *Logic, Language and Computation*, vol. 2. Stanford.

Mitchell, Peter, Robinson, Elizabeth & Thompson, D. 1999. Children's understanding that utterances emanate from minds: using speaker belief to aid interpretation. *Cognition*, 72: 45-66.

Morgan, Jerry & Green, Georgia. 1987. On the search for relevance. *Behavioral and Brain Sciences* 10: 726-7.

Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509-59.

Nicolle, Steven & Clark, Billy. 1999. Experimental pragmatics and what is said: A response to Gibbs and Moise. *Cognition* 66: 337-354.

Noh, Eun-ju. 1998. Echo questions: Metarepresentation and pragmatic enrichment. *Linguistics and Philosophy* 21: 603-628.

Noh, Eun-ju. 2001. *Metarepresentation: A Relevance-Theory Approach.* John Benjamins, Amsterdam.

Noveck, Ira. 2001. When children are more logical than adults: Investigations of scalar implicature. *Cognition* 78: 165-188.

Noveck, Ira, Bianco, Maryse & Castry, Alain. 2001. The costs and benefits of metaphor. *Metaphor and Symbol* 16: 109-121.

Origgi, Gloria and Sperber, Dan. 2000. Evolution, communication and the proper function of language. In P. Carruthers and A. Chamberlain (eds) *Evolution and the Human Mind: Language, Modularity and Social Cognition*: 140-169. Cambridge University Press, Cambridge.

Papafragou, Anna. 1998. The acquisition of modality: Implications for theories of semantic representation. *Mind and Language* 13: 370-399.

Papafragou, Anna 2000. *Modality: Issues in the Semantics–Pragmatics Interface.* Elsevier Science, Amsterdam.

Papafragou, Anna. 2002. Mindreading and verbal communication. *Mind & Language* 17: 55-67. Special Issue on Pragmatics and Cognitive Science.

Papafragou, Anna. forthcoming. Scalar implicature: Experiments at the semantics–pragmatics interface. IRCS Penn ms.

Perner, Josef, Frith, Uta, Leslie, Alan & Leekam, Sue. 1989. Explorations of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development* 60: 689-700.

Pilkington, Adrian. 2000. *Poetic Effects: A Relevance Theory Perspective.* John Benjamins, Amsterdam.

Politzer, Guy. 1990. Characterizing spontaneous inferences. *Behavioral and Brain Sciences* 13: 177-178.

Politzer, Guy & Macchi, Laura. 2000. Reasoning and pragmatics. *Mind and Society* 1: 73-93.

Premack, David & Premack, Ann James. 1994. Moral belief: Form versus content. In Laurence Hirschfeld & Susan Gelman (eds) 1994: 149-168.

Recanati, François. 1989. The pragmatics of what is said. *Mind & Language* 4: 295-329. Reprinted in Steven Davis (ed.) 1991: 97-120.

Recanati, François. 1995. The alleged priority of literal interpretation. *Cognitive Science* 19: 207-232.

Recanati, François. 2000. *Oratio Obliqua, Oratio Recta: The Semantics of Metarepresentations.* MIT Press, Cambridge MA.

Recanati, François. 2002a. Does linguistic communication rest on inference? *Mind & Language* 17: 105-126. Special Issue on Pragmatics and Cognitive Science.

Recanati, François. 2002b. Unarticulated constituents. *Linguistics and Philosophy* 25: 299-345.

Rouchota, Villy & Jucker, Andreas (eds). 1998. *Current Issues in Relevance Theory.* John Benjamins, Amsterdam.

Searle, John. 1979. Literal meaning. In J. Searle *Expression and Meaning*: 117-136. Cambridge University Press, Cambridge.

Searle, John. 1980. The background of meaning. In J. Searle, F. Keifer & M. Bierwisch (eds.) *Speech Act Theory and Pragmatics*: 221-232. Reidel, Dordrecht.

Smith, Neil & Wilson, Deirdre. 1992. Introduction to the Special Issue on Relevance theory. *Lingua* 87: 1-10.

Sperber, D. 1984 Verbal irony: pretense or echoic mention? *Journal of Experimental Psychology: General* 113: 130-6.

Sperber, Dan. 1994. Understanding verbal understanding. In J. Khalfa (ed.) *What is Intelligence?* 179-98. Cambridge University Press, Cambridge.

Sperber, Dan. 1996. *Explaining Culture: A Naturalistic Approach.* Blackwell, Oxford.

Sperber, Dan. 1997. Intuitive and reflective beliefs. *Mind and Language* 12: 67-83.

Sperber, Dan. 2000. Metarepresentations in an evolutionary perspective. In D. Sperber (ed.) *Metarepresentations: An Interdisciplinary Perspective*: 117-137. Oxford University Press, New York.

Sperber, Dan. 2002. In defense of massive modularity. In E. Dupoux (ed.) *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*: 47-57. MIT Press, Cambridge MA.

Sperber, Dan, Cara, Francisco, & Girotto, Vittorio. 1995. Relevance theory explains the selection task. *Cognition*, 57: 31-95.

Sperber, Dan & Girotto, Vittorio. forthcoming. Does the selection task detect cheater detection? To appear in *Cognition*.

Sperber, Dan & Wilson, Deirdre. 1981. Irony and the use–mention distinction. In P. Cole (ed.) *Radical Pragmatics*: 295-318. Academic Press, New York. Reprinted in Steven Davis (ed.) 1991: 550-63.

Sperber, Dan & Wilson, Deirdre. 1985/6. Loose talk. *Proceedings of the Aristotelian Society* LXXXVI: 153-71. Reprinted in Steven Davis (ed.) 1991: 540-49.

Sperber, Dan & Wilson, Deirdre. 1986a. *Relevance: Communication and Cognition.* Blackwell, Oxford and Harvard University Press, Cambridge MA. (Second edition 1995. Blackwell, Oxford.)

Sperber, Dan & Wilson, Deirdre. 1986b. On defining relevance. In R. Grandy & R. Warner (eds) *Philosophical Grounds of Rationality: Intentions, Categories, Ends*: 143-158. Oxford University Press, Oxford.

Sperber, Dan & Wilson, Deirdre. 1987a. Précis of *Relevance*. *Behavioral and Brain Sciences* 10.4: 697-710. Reprinted in Asa Kasher (ed.) 1998, vol. V: 82-115.

Sperber, Dan & Wilson, Deirdre. 1987b. Presumptions of relevance. *Behavioral and Brain Sciences* 10: 736-53.

Sperber, Dan & Wilson, Deirdre. 1990a. Spontaneous deduction and mutual knowledge. *Behavioral and Brain Sciences* 13: 179-84.

Sperber, Dan & Wilson, Deirdre. 1990b. Rhetoric and relevance. In J. Bender & D. Wellbery (eds) *The Ends of Rhetoric: History, Theory, Practice*: 140-56. Stanford University Press, Stanford CA.

Sperber, Dan & Wilson, Deirdre. 1995. Postface to the second edition of *Relevance: Communication and Cognition.* Blackwell, Oxford.

Sperber, Dan & Wilson, Deirdre. 1996. Fodor's frame problem and relevance theory: A reply to Chiappe & Kukla. *Behavioral & Brain Sciences* 19: 530-32.

Sperber, Dan & Wilson, Deirdre. 1998a. The mapping between the mental and the public lexicon. In P. Carruthers & J. Boucher (eds.) *Language and Thought: Interdisciplinary Themes*: 184-200. Cambridge University Press, Cambridge.

Sperber, Dan & Wilson, Deirdre. 1998b. Irony and relevance: A reply to Seto, Hamamoto and Yamanashi. In Robyn Carston & Seiji Uchida (eds.) 1998: 283-93.. John Benjamins, Amsterdam.

Sperber, Dan & Wilson, Deirdre. 2002. Pragmatics, modularity and mind-reading. *Mind & Language* 17: 3-23. Special Issue on Pragmatics and Cognitive Science.

Stainton, Robert. 1994. Using non-sentences: An application of relevance theory. *Pragmatics and Cognition* 2: 269-284.

Stainton, Robert. 1997. Utterance meaning and syntactic ellipsis. *Pragmatics and Cognition* 5: 51-78.

Sweetser, Eve. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge University Press, Cambridge.

Tanaka, Keiko. 1992. The pun in advertising: A pragmatic approach. *Lingua* 87: 91-102.

Traugott, Elizabeth. 1998. The role of pragmatics in semantic change. Paper presented to IPRA, Rheims, July 1998.

Travis, Charles. 1985. On what is strictly speaking true. *Canadian Journal of Philosophy* 15: 187-229.

Travis, Charles. 2000. Unshadowed Thought: Representation in Thought and Language. Harvard University Press, Cambridge MA.

van der Henst, Jean-Baptiste. 1999. The mental model theory and spatial reasoning re-examined: the role of relevance in premise order. *British Journal of Psychology* 90: 73-84.

van der Henst, Jean-Baptiste, Carles, Laure & Sperber, Dan. forthcoming. Truthfulness and relevance in telling the time. To appear in *Mind & Language*.

van der Henst, Jean-Baptiste, Sperber, Dan, & Politzer, Guy. 2002. When is a conclusion worth deriving? A relevance-based analysis of indeterminate relational problems. *Thinking & Reasoning* 8: 1-20.

van Rooy, Robert. 1999. Questioning to resolve decision problems. In P. Dekker (ed.) *Proceedings of the Twelfth Amsterdam Colloquium.* Amsterdam: ILLC.

van Rooy, Robert. 2001. Relevance of communicative acts. To appear in *Proceedings of Tark 2001*.

Wason, Peter. 1966. Reasoning. In B. M. Foss (ed.) *New Horizons in Psychology*. Penguin, Harmondsworth.

Wharton, Tim. 2001. Natural pragmatics and natural codes. *UCL Working Papers in Linguistics* 13: 109-158.

Wharton, Tim. forthcoming. Interjections, language and the 'showing'/'saying' continuum. To appear in *Pragmatics and Cognition*, January 2003.

Wharton, Tim. in preparation. *Pragmatics and the Showing–Saying Distinction*. University College London Ph.D. dissertation.

Whiten, Andrew (ed.). 1991. *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*. Oxford: Blackwell.

Wilson, Deirdre. 1998. Linguistic structure and inferential communication. In B. Caron (ed.) *Proceedings of the 16th International Congress of Linguists* (Paris, 20-25 July 1997). Elsevier Sciences, Amsterdam.

Wilson, Deirdre. 1999. Relevance and relevance theory. In R. Wilson & F. Keil (eds.) *MIT Encyclopedia of the Cognitive Sciences*: 719-22. MIT Press, Cambridge MA.

Wilson, Deirdre. 2000. Metarepresentation in linguistic communication. In D. Sperber (ed.) *Metarepresentations: An Interdisciplinary Perspective*: 411-448. Oxford University Press, Oxford.

Wilson, Deirdre & Matsui, Tomoko. 2000. Recent approaches to bridging: Truth, coherence, relevance. In J de Bustos Tovar, P. Charaudeau, J. Alconchel, S. Iglesias Recuero & C. Lopez Alonso (eds.) *Lengua, Discurso, Texto*, vol. 1: 103-132. Visor Libros, Madrid.

Wilson, Deirdre & Sperber, Dan. 1981. On Grice's theory of conversation. In P. Werth (ed.) *Conversation and Discourse.* Croom Helm, London: 155-78. Reprinted in Asa Kasher (ed.) 1998, vol. 1V: 347-68.

Wilson, Deirdre & Sperber, Dan. 1986. Pragmatics and modularity. *Chicago Linguistic Society* 22, *Parasession on Pragmatics and Grammatical Theory*: 68-74. Reprinted in Steven Davis1991: 583-95.

Wilson, Deirdre & Sperber, Deirdre. 1988. Mood and the analysis of non-declarative sentences. In J. Dancy, J. Moravcsik & C. Taylor (eds) *Human Agency: Language, Duty and Value*: 77-

101. Stanford University Press, Stanford CA. Reprinted in Asa Kasher (ed.) 1998, vol. II: 262-89

Wilson, Deirdre & Sperber, Deirdre. 1992. On verbal irony. *Lingua* 87: 53-76.

Wilson, Deirdre & Sperber, Deirdre. 1993. Linguistic form and relevance. *Lingua* 90: 1-25.

Wilson, Deirdre & Sperber, Deirdre.1998. Pragmatics and time. In Robyn Carston & Seiji Uchida (eds) 1998: 1-22.

Wilson, Deirdre & Sperber, Deirdre. 2002. Truthfulness and relevance. To appear in *Mind* 2002.

Winner, Ellen. 1988. *The Point of Words: Children's Understanding of Metaphor and Irony.* Harvard University Press, Cambridge MA.

---

* We are grateful to Larry Horn and Greg Ward for many valuable comments and suggestions on an earlier version of this paper.

[1] On the distinction between decoding and inference, see Sperber & Wilson (1986a): §1.1-5, chapter 2. On the relation between decoding and inference in comprehension, see Blakemore (1987, this volume, forthcoming); Wilson & Sperber (1993); Wilson (1998); Carston (1998, 1999, forthcoming); Origgi & Sperber (2000); Wharton (2001, forthcoming); Breheny (2002); Recanati (2002a). On the role of demonstrative and non-demonstrative inference processes in comprehension, see Sperber & Wilson (1986a): §2.1-7; Sperber & Wilson (2002); Recanati (2002a); Carston (2002, forthcoming).

[2] For early arguments against these aspects of Grice's framework, see Sperber & Wilson (1981); Wilson & Sperber (1981). For discussion and further references, see below.

[3] The notion of a POSITIVE COGNITIVE EFFECT is needed to distinguish between information that merely SEEMS to the individual to be relevant and information that actually IS relevant. We are all aware that some of our beliefs may be false (even if we cannot tell which they are), and would prefer not to waste our effort drawing false conclusions. An efficient cognitive system is one which tends to pick out genuinely relevant inputs, yielding genuinely true conclusions. For discussion, see Sperber & Wilson (1995): §3.1-2.

[4] The notion of a COGNITIVE EFFECT (or CONTEXTUAL EFFECT) has been revised several times. For early accounts, see Wilson & Sperber (1981, 1986b). For the standard definitions, see Sperber & Wilson (1986a): §2.7, and especially footnote 26. For discussion of the deductive inferences involved in deriving cognitive effects, see Politzer (1990); Sperber & Wilson (1990a). For the notion of a positive cognitive effect, see Sperber & Wilson (1995): §3.1-2. We leave open the possibility that there may be still further types of positive cognitive effect (improvements in memory or imagination, for example (cf. Wilson & Sperber 2002).

[5] For some suggestions about how this might be done, see Sperber & Wilson (1986a): 124-32. Formal notions of relevance are currently being explored by Merin (1997); Blutner (1998) (which

brings together ideas from Horn 1984, 1992; Levinson 1987; 2000; Hobbs et al. 1993; and Sperber & Wilson); van Rooy (1999, 2001). For some alternative notions of relevance, see references in Wilson & Sperber (1986b); Wilson (1999).

[6] On the distinction between COMPARATIVE and QUANTITATIVE concepts, see Carnap (1950); Sperber & Wilson (1986a): 79-81, 124-32. On comparative and quantitative notions of relevance, see Sperber & Wilson (1986a): §§3.2, 3.5, 3.6. For some factors affecting comparative assessments of relevance, see Sperber & Wilson (1986a): §§3.2, 3.6; Sperber & Wilson (1996).

[7] It is sometimes suggested that the lack of a quantitative notion of relevance makes the theory untestable. In fact, there is now a considerable experimental literature on relevance theory, and many procedures for testing and manipulating effort, effect and relevance (for discussion, see footnote 5 and §6 below.)

[8] This is the simpler of two characterisations of ostensive-inferential communication in Sperber & Wilson (1986a): 29, 58, 61. The fuller characterisation involves the notions of MANIFESTNESS and MUTUAL MANIFESTNESS. In particular, we argue that for communication to be truly overt, the communicator's informative intention must become not merely manifest to the audience (i.e. capable of being recognised and accepted as true, or probably true), but mutually manifest to communicator and audience. On the communicative and informative intentions, see Sperber & Wilson (1986a): §1.9-12; on the notion of mutual manifestness, see Garnham & Perner (1990); Sperber & Wilson (1990a).

[9] For arguments against the view that co-operation in Grice's sense is fundamental to communication, see Wilson & Sperber (1981); Sperber & Wilson (1986a): 161-2; Smith & Wilson (1992); Sperber (1994). For more general arguments that rationality in communication does not require co-operation in Grice's sense, see Kasher (1976); Sperber (2000); Sperber & Wilson (2002).

[10] On the use of silence as an ostensive stimulus, see Morgan & Green (1987): 727; Sperber & Wilson (1987b): 746-7.

[11] The analysis of scalar implicatures is another case where Gricean analyses tend to lose the symmetry between unwillingness and inability to provide relevant information. For discussion, see Sperber & Wilson (1995): 276-8; Green (1995); Matsumoto (1995); Carston (1995, 1998a); and §6 below. For experimental work, see Noveck (2001); Papafragou (2002, forthcoming).

[12] For discussion and illustration, see Carston (this volume). On the notion of explicit content, see §4 below.

[13] Notice, incidentally, that the hearer's expectations of relevance may be readjusted in the course of comprehension. For example, it may turn out that the effort of finding any interpretation at all would be too great: as a result, the hearer would disbelieve the presumption of relevance and terminate the process, with his now null expectations of relevance trivially satisfied.

[14] It is sometimes suggested (e.g. by Morgan and Green 1987: 726-7) that puns and deliberate equivocations present a problem for this approach. We would analyse these as cases of layering

in communication, a widespread phenomenon which fits straightforwardly with our account. Just as the failure to provide relevant information at one level may be used as an ostensive stimulus at another, so the production of an utterance which is apparently uninterpretable at one level may be used as an ostensive stimulus at another (see Sperber & Wilson 1987b: 751; Tanaka 1992).

[15] By 'explicitly communicated content' (or EXPLICATURE), we mean a proposition recovered by a combination of decoding and inference, which provides a premise for the derivation of contextual implications and other cognitive effects (Sperber & Wilson 1986a: 176-93; Carston this volume; forthcoming). Despite many terminological disagreements (see footnotes 23 and 24), the existence of pragmatic contributions at this level is now widely recognised (see e.g. Wilson & Sperber 1981, 1998, 2002; Kempson & Cormack 1982; Travis 1985, 2001; Sperber & Wilson 1986a: §4.2-3; Kempson 1986, 1996; Blakemore 1987; Carston 1988, 2000, 2002, forthcoming; Recanati 1989, 2002b; Neale 1992; Bach 1994a, 1994b, 1997; Stainton 1994, 1997, this volume; Bezuidenhout 1997; Levinson 2000; Fodor 2001).

[16] In his 'Retrospective Epilogue', and occasionally elsewhere, Grice seems to acknowledge the possibility of intentional pragmatic contributions to 'dictive content' (Grice 1989: 359-68). See Carston (forthcoming); Wharton (in preparation) for discussion.

[17] On the distinction between primary and secondary pragmatic processes, see Breheny (2002); Recanati (2002a); Carston (this volume, forthcoming); Sperber & Wilson (2002). Some of the literature on generalised conversational implicature and discourse pragmatics tacitly invokes a similar distinction (cf. Hobbs 1985; Lascarides & Asher 1993; Lascarides, Copestake & Briscoe 1996; Levinson 2000). See also footnotes 23 and 24.

[18] See, for example, Sperber & Wilson (1986a): §4.3-5, esp. pp 204-208; Wilson & Sperber (2002).

[19] A hearer's expectations of relevance may be more or less sophisticated. In an unsophisticated version, presumably the one always used by young children, what is expected is actual optimal relevance. In a more sophisticated version (used by competent adult communicators who are aware that the speaker may be mistaken about what is relevant to the hearer, or in bad faith and merely intending to appear relevant), what is expected may be merely attempted or purported optimal relevance. Adult communicators may nevertheless expect actual optimal relevance by default. Here we will ignore these complexities, but see Sperber (1994); Wilson (2000); and §5 below.

[20] For expository purposes, we have chosen an example in which the linguistic content of the discourse, and in particular the preceding utterance ("No"), creates a fairly precise expectation of relevance, allowing the interpretation process to be strongly driven by expectations of effect. In an indirect answer such as (ib), where the linguistic form of the utterance is compatible with two different lines of interpretation, considerations of effort, and in particular the accessibility of contextual assumptions capable of yielding the expected conclusions, play a more important role. In a discourse-initial utterance such as (ii), or in a questionnaire situation, considerations of effort are likely to play a decisive role in narrowing down the possible lines of interpretation:

(i) a. *Peter*: Did John pay back the money he owed?

b. *Mary:* He forgot to go to the bank.

(ii) He forgot to go to the bank.

[21] For one thing, we have used English sentences to represent the assumptions and assumption schemas that Peter entertains at different stages of the comprehension process, which we assume he does not represent in English but in some conceptual representation system or language of thought. We have also left aside semantic issues such as the analysis of the definite article and definite descriptions (e.g. *the bank*).

[22] See for example Horn (1984, 1992); Levinson (1987, 2000); Hobbs et al. (1993); Lascarides, Copestake & Briscoe (1996); Lascarides & Copestake (1998); Blutner (1998, 2002).

[23] As noted above (footnote 15), there is some debate about how the explicit–implicit distinction should be drawn (see, for example, Horn 1992; Sperber & Wilson 1986a: §4.1-4; Wilson & Sperber 1993; Bach 1994a,b, 1997; Levinson 2000; Carston 2002, this volume, forthcoming). The issue is partly terminological, but becomes substantive when combined with the claim that explicit and implicit communication involve distinct pragmatic processes (as it is in much of the literature on generalised implicatures: e.g. Levinson 2000).

[24] Levinson (2000: 195-6) discusses a number of possible criteria for distinguishing explicatures from implicatures, provides arguments against each, and concludes that the distinction is unjustified. But there is no reason to expect a criterion to be provided for each theoretical distinction. (We would not expect the defenders of a distinction between generalised and particularised implicatures to provide a criterion, although we would expect them to characterise these notions clearly and provide sound supporting evidence.) Our notion of an explicature is motivated, among other things, by embedding tests which suggest that certain pragmatic processes contribute to truth-conditional content, while others do not (Wilson & Sperber 1986a: 80; 2002). The allocation of pragmatically inferred material between explicatures and implicatures is constrained, on the one hand, by our theoretical definitions of explicature and implicature (Sperber & Wilson 1986a: 182), and, on the other, by the fact that the implicated conclusions which satisfy the hearer's expectations of relevance must be warranted by the explicit content of the utterance, together with the context. For further discussion, see Sperber & Wilson (1986a): §4.3; Sperber & Wilson (1998a); Carston (1995, 1998, 2000, this volume); Wilson & Sperber (1998, 2002). For some experimental work, see Gibbs & Moise (1997); Matsui (1998, 2000); Nicolle & Clark (1999); Wilson & Matsui (2000); Noveck (2001); Papafragou (2002, forthcoming).

[25] Since lexical loosening is widely acknowledged as one of the factors driving semantic change, it might be argued that from a synchronic point of view, these are simply cases of polysemy. However, we are interested in the pragmatic micro-processes underlying these semantic changes, and we will largely abstract away from the question of whether *Hoover*, or *square*, or *silent* has acquired an extra stable sense. Notice, though, that the variation in interpretations of a word such as *square* or *silent* applied to different objects in different circumstances is so great as to make purely semantic or default-pragmatic explanations seem unfeasible (for discussion, see Searle 1979, 1980; Horn 1984; Lakoff 1987; Franks & Braisby 1990; Sweetser 1990; Hobbs et al. 1993; Bach 1994a,b, 1997; Recanati 1995; Carston 1997, 1998, this volume, forthcoming; Sperber &

Wilson 1998a; Traugott 1998; Wilson 1998; Lasersohn 1999; Asher & Lascarides 2001; Papafragou 2000; Wilson & Sperber 2002).

[26] For early arguments against the maxim of truthfulness, see Wilson & Sperber (1981). For detailed critiques of frameworks based on maxims or conventions of truthfulness, discussion of some existing accounts of loose use, and justification of an alternative, relevance-theoretic account, see Wilson & Sperber (2002). For experimental evidence, see Matsui (1998, 2000); Wilson & Matsui (2000); van der Henst, Carles & Sperber (forthcoming).

[27] For Grice, metaphor and hyperbole involve different types of interpretation process, and may indeed be mutually exclusive: see Grice (1989): 34.

[28] See, for example, Gibbs (1994); Noveck, Bianco & Castry (2001); Glucksberg (2001). Glucksberg's view that the interpretation of metaphor involves the construction of a broader category than the one determined by the encoded meaning fits well with our analysis based on loose use.

[29] While the claim that metaphor is a variety of loose use has been part of the theory for some time (see e.g. Sperber & Wilson 1985/6, 1986a, §4.7-8, 1990b), the details of this analysis are new. For discussion, see Recanati (1995); Carston (1997, this volume, forthcoming); Sperber & Wilson (1998a); Wilson & Sperber (2002).

[30] Or an appropriately narrowed-and-loosened variant.

[31] On the notion of interpretive use, see Sperber & Wilson (1986a): §4.7; Blass (1990); Gutt (1991); Sperber (1997); Wilson (2000); Noh (2001); Papafragou (1998, 2000). On the notion of echoic use, see Sperber & Wilson (1986a): §4.9; Blakemore (1994); Carston (1996, forthcoming); Noh (1998); Wilson (2000).

[32] The relevance-theoretic account of irony was first proposed in Sperber & Wilson 1981. It was extended and developed in Sperber & Wilson (1986a): §§4.7, 4.9; Sperber & Wilson (1990b, 1998b); Wilson & Sperber (1992); Curcò (1998). For critical discussion, see Clark & Gerrig (1984); Kreuz & Glucksberg (1989); Gibbs & O'Brien (1992); Martin (1992); Kumon-Nakamura, Glucksberg & Brown (1995); and the papers by Seto, Hamamoto and Yamanashi in Carston & Uchida, eds. (1998). For responses, see Sperber (1984); Sperber & Wilson (1998b).

[33] This approach has been experimentally tested: see Jorgensen, Miller & Sperber (1984); Happé (1993); Gibbs (1994); Kreuz & Glucksberg (1989); Gibbs & O'Brien 1992; Kumon-Nakamura, Glucksberg & Brown (1995); Langdon, Davies & Coltheart (2002).

[34] On the development of metaphor and irony, see Winner (1988). On the relation between irony, metaphor and metarepresentational abilities, see Happé (1993); Langdon, Davies & Coltheart (2002). For further discussion of the relation between communicative and metarepresentational abilities, see §5 below.

[35] Levinson (2000: 239) interprets us (mistakenly) as claiming that ironies "are implicatures interpreted as 'echoes' of what someone might have said: they are distinctly not explicatures". He

objects that our account does not allow for the fact that ironical use of a referential expression may make a difference to truth conditions (as in his nice example "If you need a car, you may borrow my Porsche" [used to refer to the speaker's VW]). In fact, such examples provide strong confirmation of our account, on which irony is closely related to mention, quotation and other types of metalinguistic use, and hence contributes directly to explicatures. It is well known that metalinguistic use of a word may make a difference to truth conditions (see Horn 1989; Sperber & Wilson 1981; 1986a: §4.7; Carston 1996, forthcoming; Cappelen & Lepore 1997; Noh 2000; Wilson 2000.)

[36] For discussion, see Sperber & Wilson (1986a): §1.9-12.

[37] In the relevance-theoretic framework, mood indicators are among the items seen as carrying procedural rather than conceptual meaning. For discussion, see Blakemore (1987, this volume, forthcoming); Wharton (forthcoming, in preparation).

[38] On higher-level explicatures, see Blakemore (1991); Wilson & Sperber (1993); Ifantidou (2001). On the analysis of non-declarative utterances, see Sperber & Wilson (1986a): §4.10; Wilson & Sperber (1988); Wilson (2000); Noh (2001). For critical discussion, see Bird (1994); Harnish (1994).

[39] See, for example, Whiten (1991); Davies and Stone (1995a,b); Carruthers & Smith (1996); Malle, Moses & Baldwin (2001).

[40] See, for example, Perner, Frith, Leslie & Leekam (1989); Happé (1993); Baron-Cohen (1995); Mitchell, Robinson & Thompson (1999); Happé & Loth (2002); Papafragou (2002); and the papers in *Mind & Language* 17.1-2 (2002).

[41] We are using "module" in a somewhat looser sense than Fodor's, to mean a domain- or task-specific autonomous computational mechanism (for discussion, see Sperber 1996: chapter 6).

[42] See, for example, Leslie (1991); Hirschfeld & Gelman (1994); Barkow, Cosmides & Tooby (1995); Sperber (1996, 2002). For critical comments, see Fodor (2000).

[43] For explicit defence of this position, see Bloom (2000, 2002). For experimental evidence in favour of a more modular approach, see Happé & Loth (2002).

[44] For other applications of relevance theory in the experimental study of reasoning, see Politzer & Macchi (2000), van der Henst (1999) van der Henst, Sperber, & Politzer (2002).